

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/641,808
Source: FW/6
Date Processed by STIC: 1/21/05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 01/21/2005

PATENT APPLICATION: US/09/641,808

TIME: 15:36:35

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw

```

3 <110> APPLICANT: BELFORT, MARLENE
4     BELFORT, GEORGES
5     DERBYSHIRE, VICKY
6     WOOD, DAVID
7     WU, WEI
9 <120> TITLE OF INVENTION: GENETIC SYSTEM AND SELF-CLEAVING INTEINS DERIVED
10    THEREFROM, BIOSEPARATIONS AND PROTEIN PURIFICATION
11    EMPLOYING SAME, AND METHODS FOR DETERMINING CRITICAL,
12    GENERALIZABLE AMINO ACID RESIDUES FOR VARYING
14 <130> FILE REFERENCE: 454311-2201.1
16 <140> CURRENT APPLICATION NUMBER: 09/641,808
17 <141> CURRENT FILING DATE: 2000-08-17
19 <150> PRIOR APPLICATION NUMBER: 60/149,257
20 <151> PRIOR FILING DATE: 1999-08-17
22 <160> NUMBER OF SEQ ID NOS: 26
24 <170> SOFTWARE: PatentIn Ver. 3.3
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 7
28 <212> TYPE: PRT
29 <213> ORGANISM: Artificial Sequence
31 <220> FEATURE:
32 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
33    peptide
35 <400> SEQUENCE: 1
36 Cys Gly Glu Gln Pro Thr Gly
37 1           5
40 <210> SEQ ID NO: 2
41 <211> LENGTH: 5
42 <212> TYPE: PRT
43 <213> ORGANISM: Artificial Sequence
45 <220> FEATURE:
46 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
47    peptide
49 <400> SEQUENCE: 2
50 Ser Ile Glu Gln Asp
51 1           5
54 <210> SEQ ID NO: 3
55 <211> LENGTH: 5
56 <212> TYPE: PRT
57 <213> ORGANISM: Artificial Sequence
59 <220> FEATURE:
60 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
61    peptide

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/641,808

DATE: 01/21/2005

TIME: 15:36:35

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw

```

63 <400> SEQUENCE: 3
64 Cys Arg Ala Met Gly
65   1           5
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 5
70 <212> TYPE: PRT
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
75     peptide
77 <400> SEQUENCE: 4
78 Met Ile Glu Gln Asp
79   1           5
82 <210> SEQ ID NO: 5
83 <211> LENGTH: 19
84 <212> TYPE: PRT
85 <213> ORGANISM: Foot-and-mouth disease virus
87 <400> SEQUENCE: 5
88 Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
89   1           5             10             15
91 Pro Gly Pro
95 <210> SEQ ID NO: 6
96 <211> LENGTH: 8
97 <212> TYPE: DNA
98 <213> ORGANISM: Artificial Sequence
100 <220> FEATURE:
101 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
102     polynucleotide sequence
104 <400> SEQUENCE: 6
105 agtcagtc                                     8
108 <210> SEQ ID NO: 7
109 <211> LENGTH: 8
110 <212> TYPE: DNA
111 <213> ORGANISM: Artificial Sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: Description of Artificial Sequence: Illustrative
115     polynucleotide sequence
117 <400> SEQUENCE: 7
118 aatcaatc                                     8
121 <210> SEQ ID NO: 8
122 <211> LENGTH: 19
123 <212> TYPE: PRT
124 <213> ORGANISM: Foot-and-mouth disease virus
126 <400> SEQUENCE: 8
127 Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
128   1           5             10             15
130 Pro Gly Pro
134 <210> SEQ ID NO: 9
135 <211> LENGTH: 9

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/641,808

DATE: 01/21/2005

TIME: 15:36:35

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw

136 <212> TYPE: PRT
 137 <213> ORGANISM: Saccharomyces cerevisiae
 139 <400> SEQUENCE: 9
 140 Phe Thr Cys Asn Ala Thr His Glu Leu
 141 1 5
 144 <210> SEQ ID NO: 10
 145 <211> LENGTH: 9
 146 <212> TYPE: PRT
 147 <213> ORGANISM: Mycobacterium leprae
 149 <400> SEQUENCE: 10
 150 Phe Ala Ala Thr Pro Asn His Leu Ile
 151 1 5
 154 <210> SEQ ID NO: 11
 155 <211> LENGTH: 9
 156 <212> TYPE: PRT
 157 <213> ORGANISM: Mycobacterium xenopi
 159 <400> SEQUENCE: 11
 160 Val Thr Gly Thr Ala Asn His Pro Leu
 161 1 5
 164 <210> SEQ ID NO: 12
 165 <211> LENGTH: 9
 166 <212> TYPE: PRT
 167 <213> ORGANISM: Pyrococcus sp.
 169 <400> SEQUENCE: 12
 170 Ile Thr Ile Thr Glu Gly His Ser Leu
 171 1 5
 174 <210> SEQ ID NO: 13
 175 <211> LENGTH: 9
 176 <212> TYPE: PRT
 177 <213> ORGANISM: Porphyra purpurea
 179 <400> SEQUENCE: 13
 180 Leu Glu Leu Thr Ser Asn His Lys Ile
 181 1 5
 184 <210> SEQ ID NO: 14
 185 <211> LENGTH: 9
 186 <212> TYPE: PRT
 187 <213> ORGANISM: Mycobacterium tuberculosis
 189 <400> SEQUENCE: 14
 190 Val Trp Ala Thr Pro Asp His Lys Val
 191 1 5
 194 <210> SEQ ID NO: 15
 195 <211> LENGTH: 9
 196 <212> TYPE: PRT
 197 <213> ORGANISM: Caenorhabditis elegans
 199 <400> SEQUENCE: 15
 200 Leu Arg Ile Thr Ser Arg His Phe Met
 201 1 5
 204 <210> SEQ ID NO: 16
 205 <211> LENGTH: 9

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/641,808

DATE: 01/21/2005

TIME: 15:36:35

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw

206 <212> TYPE: PRT
 207 <213> ORGANISM: Drosophila melanogaster
 209 <400> SEQUENCE: 16
 210 Leu Thr Val Thr Pro Ala His Leu Val
 211 1 5
 214 <210> SEQ ID NO: 17
 215 <211> LENGTH: 9
 216 <212> TYPE: PRT
 217 <213> ORGANISM: Xenopus laevis
 219 <400> SEQUENCE: 17
 220 Ile Arg Leu Thr Ala Ala His Leu Leu
 221 1 5
 224 <210> SEQ ID NO: 18
 225 <211> LENGTH: 19
 226 <212> TYPE: PRT
 227 <213> ORGANISM: Saccharomyces cerevisiae
 229 <400> SEQUENCE: 18
 230 Asp Tyr Tyr Gly Ile Thr Leu Ser Asp Asp Ser Asp His Gln Phe Leu
 231 1 5 10 15
 233 Leu Ala Asn
 237 <210> SEQ ID NO: 19
 238 <211> LENGTH: 18
 239 <212> TYPE: PRT
 240 <213> ORGANISM: Mycobacterium leprae
 242 <400> SEQUENCE: 19
 243 Ser Met Asn Arg Phe Asp Ile Glu Val Glu Gly Asn His Asn Tyr Phe
 244 1 5 10 15
 246 Val Asp
 250 <210> SEQ ID NO: 20
 251 <211> LENGTH: 19
 252 <212> TYPE: PRT
 253 <213> ORGANISM: Mycobacterium xenopi
 255 <400> SEQUENCE: 20
 256 Val Gln Pro Val Tyr Ser Leu Arg Val Asp Thr Ala Asp His Ala Phe
 257 1 5 10 15
 259 Ile Thr Asn
 263 <210> SEQ ID NO: 21
 264 <211> LENGTH: 18
 265 <212> TYPE: PRT
 266 <213> ORGANISM: Pyrococcus sp.
 268 <400> SEQUENCE: 21
 269 Asp Gly Tyr Val Tyr Asp Leu Ser Val Asp Glu Asp Glu Asn Phe Leu
 270 1 5 10 15
 272 Ala Gly
 276 <210> SEQ ID NO: 22
 277 <211> LENGTH: 18
 278 <212> TYPE: PRT
 279 <213> ORGANISM: Porphyra purpurea
 281 <400> SEQUENCE: 22

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/641,808

DATE: 01/21/2005

TIME: 15:36:35

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw

```

282 Phe Gln Asn Val Phe Asp Phe Ala Ala Asn Pro Ile Pro Asn Phe Ile
283   1           5           10           15
285 Ala Asn
289 <210> SEQ ID NO: 23
290 <211> LENGTH: 18
291 <212> TYPE: PRT
292 <213> ORGANISM: Mycobacterium tuberculosis
294 <400> SEQUENCE: 23
295 Arg Ala Arg Thr Phe Asp Leu Glu Val Glu Glu Leu His Thr Leu Val
296   1           5           10           15
298 Ala Glu
302 <210> SEQ ID NO: 24
303 <211> LENGTH: 16
304 <212> TYPE: PRT
305 <213> ORGANISM: Caenorhabditis elegans
307 <400> SEQUENCE: 24
308 Thr Gly Ile Tyr Ser Pro Leu Thr Asn Asn Gly Arg Ile Ile Val Asn
309   1           5           10           15
312 <210> SEQ ID NO: 25
313 <211> LENGTH: 16
314 <212> TYPE: PRT
315 <213> ORGANISM: Drosophila melanogaster
317 <400> SEQUENCE: 25
318 Lys Gly Val Val Ala Pro Leu Thr Arg Glu Gly Thr Ile Val Val Asn
319   1           5           10           15
322 <210> SEQ ID NO: 26
323 <211> LENGTH: 16
324 <212> TYPE: PRT
325 <213> ORGANISM: Xenopus laevis
327 <400> SEQUENCE: 26
328 Thr Gly Ala Tyr Ala Pro Leu Thr Ala His Gly Thr Val Val Ile Asp
329   1           5           10           15

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/641,808

DATE: 01/21/2005

TIME: 15:36:36

Input Set : A:\45312201.app

Output Set: N:\CRF4\01212005\I641808.raw